DOCUMENT RESUME

AL 001 142

ED 024 015

By-Roberts, A. Hood

International Information Flow in Linguistics.

Center for Applied Linguistics, Washington, D.C.
Pub Date Dec 68
Note-3p.; Revision of paper presented at U.S.-Japan Seminar on Computational Linguistics, Honolulu, Hawaii, March 25-27, 1968.
Journal Cit-The Linguistic Reporter; v10 n6 Dec 1968
EDRS Price MF-$0.25 HC-$0.25
Descriptors-Communication Problems, Foreign Language Periodicals, Information Centers, *Information Dissemination, Information Retrieval, *Information Systems, *Interdisciplinary Approach, *International Programs, Language Research, *Linguistics, Manpower Needs, Periodicals
Identifiers-*Information Flow

Since 1967 the Center for Applied Linguistics has been conducting a survey and analysis of the field of the language sciences to determine the operational and technical requirements for an information system for the field. The first important fact to emerge from this study is that linguistics is no longer a very clearly definable area. It has become a tool widely used in areas not traditionally its concern and by people who are not linguists by training. This fact, plus the growing number of linguists and language researchers, complicates the problem of information flow because there is very little coordination of their efforts. The author discusses these complications, among which are: (1) at present there are no adequate central repositories for even the 400 core publications, (2) many periodicals have serious publication lags, (3) linguistic literature is published in a variety of languages, (4) pertinent information cuts across several fields and retrieval is therefore difficult, (5) there is no central abstracts journal, (6) there are only three or four dictionaries which cover the field adequately, and (7) there are no adequate classifications of the field. Also discussed are improvements in the area of information flow: (1) at least 63 information services in the language sciences now exist around the world, (2) various bibliographies have been published by CAL, and (3) several rosters of linguists are being maintained. (DO)

The Linguistic Reporter

Volume 10 Number 6 December 1968

Newsletter of the Center for Applied Linguistics
1717 Massachusetts Avenue, N.W.
Washington, D. C. 20036

ED 024013

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

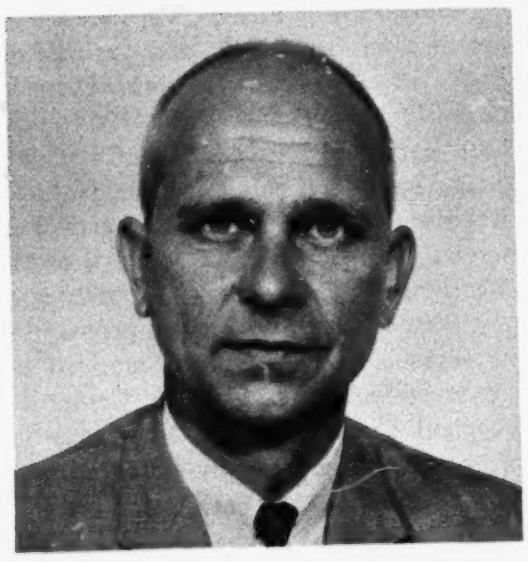
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

International Information Flow in Linguistics

by A. Hood Roberts

[A. Hood Roberts is Associate Director of the Center for Applied Linguistics. This paper was originally presented at the meeting of the U.S.-Japan Seminar on Computational Linguistics, held in Honolulu, Hawaii, March 25–27, 1968. It is printed here with minor changes by the author.]

In April 1964, the *Linguistic Reporter* published an article by Charles A. Ferguson entitled "Information Flow in Linguistics." In the article, Dr. Ferguson discussed the characteristics of information in linguistics, the needs of the field and how these needs were being met at the time. He also mentioned steps then under way to improve the dissemination of information. Since that time, the Center for Applied Linguistics has begun a survey and analysis of the field of the language sciences to determine the operational and technical requirements needed for an information system for the field. The project, known as LINCS (Language Information Network and Clearinghouse System), which began in June 1967 with National Science Foundation funding, has gathered a considerable amount of data on the current situation with respect to international information flow in linguistics. Since 1964, the situation has improved in some areas and changed in others. Some problems still remain to be solved. The first important fact to emerge from this study is that linguistics is no longer a very clearly definable area. Like mathematics, it has become a tool widely used in areas not traditionally its concern and by people who are not linguists by training. Two sets of data in particular support this assertion: (1) information collected on groups engaged in language research or providing information services in the language sciences, and (2) data collected on the periodical literature of the language sciences. This fact, plus the growing number of linguists and language researchers, complicates the problem of information flow. While the number of workers in the field is increasing, there is, unfortunately, very little coordination of their activities. As an example, only a few years ago, two translations of Leonard Bloomfield's *Language* were done independently in Latin America. The two translators learned of each other's work only when one of them published his translation ahead of the other. If the efforts of linguists were coordinated, their growing number would give strength to the field, whereas now the increase is merely aggravating the problems of information flow.



A. Hood Roberts

1. *Information services in the language sciences.* A preliminary inventory of such services, compiled by the Center for Applied Linguistics, lists 63 groups around the world which store and retrieve information of concern to the language sciences and provide publicly available services. The list includes several types of groups: (a) Some groups are concerned with area studies, e.g. the *International African Institute,* whose linguistic information work is carried out as part of its general information services in the field of African cultural and social studies. (b) Other groups are concerned with gathering material on a specific language and its dialects, or on a specific group of languages, e.g. the *Deutsches Spracharchiv,* which has a collection of tape recordings of German dialects and of the German vernacular and literary language. (c) Some other organizations are concerned with specific subfields of activity, e.g. the *Centre for Information on Language Teaching,* which collects, coordinates and disseminates information about all aspects of modern languages and their teaching. (d) There are a few centers which collect information on the language of their country for the purpose of supplying information needed by the government in connection with language

See Information Flow, 2, Col. 1

AL 001 142

INFORMATION FLOW—from page 1

planning and policy. One such organization is the *National Language Research Institute,* whose clientele includes the Language Committee of the Ministry of Education of Japan and other government offices. The *Bureau of Ghana Languages* is in the same category, having been established by the government to provide materials to support its mass literacy drive. (e) Special libraries, such as the *Newberry Library,* whose collections contain substantial language materials, are also included.

2. *Periodical literature.* The existence of such information centers will, it is hoped, provide part of the solution to the difficulties of coping with the vast periodical literature, which has to be reduced to manageable proportions. The Center for Applied Linguistics has so far identified approximately 2,000 periodicals throughout the world which deal, if only in part, with the language sciences. If we define core linguistics journals as those which publish articles written by people trained as linguists for other linguists in the field, and whose editors and editorial staff are composed of linguists, then the total number of core journals is about 400. If to that number we add the journals which handle closely related material plus the journals which, although concerned with other or larger fields, regularly publish a substantial amount of linguistic literature, the figure does not exceed 600. The rest of the literature is scattered in periodicals all over the world and there is no way at present for a linguist to be certain that he has covered his specialty adequately. There are still no adequate central repositories, even for the core material. The Library of Congress, for example, does not cover the Far Eastern material adequately, and, in fact, has dropped some of the North Korean periodicals. Even if the individual researcher knows of a particular journal and succeeds in finding a library which carries it, he may find that the most recent issue available to him is two years out-of-date, sometimes because of a delay in acquisition but often, in the case of some major sources, e.g. *Language* or *Mechanical Translation,* because of a publication lag. The March 1967 issue of *Language* did not appear until early 1968. The December 1965 issue of *Word* appeared only after a two-year delay. *PADS* (Publication of the American Dialect Society), No. 45, for April 1966, came out almost two years late, as did the preceding issues. *Mechanical Translation* has not appeared since mid-1967. Another problem the linguist has is a problem of translation. The literature of linguistics is published in a wide variety of languages. This is not the case at all in such fields as High-Particle Physics, where the papers are published in two or three major languages. Even if the scholar reads several languages other than his own, there will probably be certain languages of importance in the publishing pattern of linguistics that he will not be able to read. For speakers of other Indo-European languages, for example, Russian is not as inaccessible as it used to be, but very few Western scholars can read Japanese; they must rely on those Japanese journals that publish part of their material in languages other than Japanese or that provide abstracts in other languages. *Studia Phonologica,* published by the Institution of Phonetic Sciences (University of Kyoto), publishes articles in English, German, and Japanese. This is a regular pattern in countries whose scientists fear that their work will remain unnoticed internationally if not published in languages of wider communication. The *Norsk Tidsskrift for Sprogvindenskap* also publishes in English and other languages. The *Revue roumaine de linguistique* publishes in English, French, German, and Russian, a publication pattern which is becoming increasingly common. Unfortunately, none of these efforts to make their material more accessible are very well coordinated, and, in the case of abstracts in other languages, the scientist who wants to read the original article still has to have it translated if he is unable to read the language of the original.

One of the most difficult problems results from the increasingly interdisciplinary nature of the language science field. The researcher in psycholinguistics who is interested in aphasia, where the major breakthroughs occurred as a result of the large volume of cases of brain trauma in World War II veterans, will have to track down articles in a vast mass of literature cutting across a number of fields. It is true that he can search the literature in *Index Medicus,* but the retrieval capacity of MEDLARS is still not very good, and the behavioral science field is still not adequately covered or indexed. There is a real need for coordination at the international level of these problems of coverage. One approach to this would be to have the field divided in such a way that an individual language would be covered by a single center which would publish the material on a given language written in that language. If the center had sufficient facilities, it could process material on one language written in any language. Another solution would be to publish material in a particular subfield. A combination of these activities and coordination in the publication of secondary sources would go a long way to alleviate the problem.

There are currently no more than about forty abstracting or bibliographical publications which process a substantial amount of material relevant to the language sciences. There is no central abstracts journal to serve the community at large, and the *Linguistic Bibliography* has a serious publishing lag. There are some current awareness publications and services. The *Language Research In Progress* project of the Center for Applied Linguistics issues regular listings of ongoing research projects. The ERIC Clearinghouse for Linguistics, located at CAL, has initiated a program known as PEGS (Program for Exchange of Generative Studies), which provides for rapid dissemination of unreferreed papers in the field of generative grammar. *Dissertation Abstracts* provides abstracts only for dissertations written at the participating universities, and there is almost no way of locating master's theses. *Language and Language Behavior Abstracts* covers psycholinguistics, and *The Finite String,* newsletter of the Association for Computational Linguistics, regularly publishes a bibliography of computational linguistics, formerly supplied by the Rand Corporation and now prepared by the Linguistic Documentation Program of CAL.

3. *Basic tools.* Since 1964, *Information Sources in Linguistics,* a selective bibliographical listing of the most useful publications in linguistics, and *A Bibliography of American Doctoral Dissertations in Linguistics: 1900–1964* have been published by CAL. The list of language names prepared for the National Register of Scientific and Technical Personnel has reached 17,000. Work on the *Multilingual Thesaurus of Languages of the World,* which was under preparation at Indiana University, was stopped because of lack of funds. The Center for Applied Linguistics has taken over the slips, which are presently stored at CAL, but no decision has been made on what to do with them. A new translation journal has appeared, *Automatic Documenta-*

tion and Mathematical Linguistics, which will publish selected translations from *Nauchno-Tekhnicheskaya Informatsiya;* however, the subscription price is very high and it is unlikely that many individuals can afford to subscribe to it. There are currently about eighty dictionaries of terminology for linguistics as a whole or for linguistics with reference to a single language. However, there are only three or four which adequately cover the field. The work on the Gleason dictionary of linguistic terminology, which was being carried out at the Hartford Seminary Foundation, has been halted for lack of manpower and funds. As far as classification systems are concerned, there are still no adequate classifications of the field.

The Center for Applied Linguistics has examined a number of thesauri which include terms of relevance to the field of language sciences. These include the Educational Resources Information Center (ERIC) Thesaurus, the Western Reserve Thesaurus developed for the U.S. Office of Education, the Project LEX Thesaurus of the Department of Defense, the Thesaurus of the Johns Hopkins Information Center for Hearing, Speech and the Disorders of Human Communication, and several smaller lists. None of these adequately cover the language sciences, but some 2,500 relevant terms have been extracted from them and merged in an alphabetical list which shows the broad, narrow and related term structuring. These, however, mainly cover only the interfaces of linguistics. The ERIC thesaurus has some core terms, but these are only now being added and so far the number is small. In addition, ERIC does not cover historical linguistics. The need is for a thesaurus which would adequately cover the whole field and which might also serve as a means of achieving some standardization in the usage of technical terminology. The situation in linguistics is particularly bad in this respect, with different sets of terms for different schools. It is hoped that the initial list can be used as the basis for a thesaurus which will, after consultation with specialists, provide both a tool for efficient storage and retrieval and a guide to the usage of preferred terms.

4. *Manpower.* The National Science Foundation, under the National Science Foundation Act of 1950, maintains a National Register of Scientific and Technical Personnel and provides a clearinghouse for information covering all scientific and technical personnel in the United States. Approximately 250,000 scientists responded to the questionnaire circulated in 1966, which contains items on educational background, current employment, specialization in science or technology, and selected personal characteristics. The questionnaire also contains items on language competence and linguistic specialties. The questionnaires are distributed by major professional societies; those for linguistics are distributed through CAL. In 1966, 1,269 linguists returned the forms. The data derived from the questionnaire are used in the preparation of statistical studies of the U.S. linguistic community. Additionally, CAL maintains a Roster of Linguists, a manpower file containing a total of approximately 4,000 names; the Linguistic Research Group of Pakistan has published a *Directory of Pakistani Linguists and Language Scholars;* the Bureau pour l'Enseignement de la Langue et de la Civilisation françaises maintains a file on French linguists and other linguists who are working in France; and the National Language Research Institute in Japan has a "List of Persons" containing the names of 1,500 persons in charge of the study of the Japanese language, language problems, and language education. The Inventory of Information Services in the Language Sciences, a project of CAL, contains information on other groups which maintain manpower files of this sort.

5. *Informal communication.* No adequate studies of communication patterns exist at this time. It is only possible to say generally that the new interdisciplinary atmosphere has led to fruitful personal exchanges. Linguists are now beginning to attend large international scientific conferences, and it is hoped that this will pave the way for cooperation on an international scale.

The Center for Applied Linguistics is a nonprofit, internationally oriented professional organization, established in 1959 and incorporated in 1964 in Washington, D.C. The purpose of the Center is to serve as a clearinghouse and informal coordinating body in the application of linguistics to practical language problems.

The Linguistic Reporter, the Center's newsletter, is published six times a year, in February, April, June, August, October, and December. Annual subscription, $1.50; air mail, $3.50. (Individuals faced with currency restrictions or similar limitations are invited to write to the Editor.) Manuscripts, books for review, and editorial communications should be sent to Frank A. Rice, Editor, THE LINGUISTIC REPORTER, Center for Applied Linguistics, 1717 Massachusetts Avenue, N.W., Washington, D.C. 20036. Communications concerning subscriptions should be directed to the Subscriptions Secretary at the same address. Permission is granted for quotation or reproduction from the contents of the LINGUISTIC REPORTER provided acknowledgment is given.